Business Research Methods: Measurement and Data collection

By Dr. Satyabrata Dash

Professor- MBA Marketing

SMIT- PGCMS, Brahmapur

Primary and Secondary Data

- The primary data are those which are collected afresh and for the first time, and thus happen to be original in character. Such data are published by authorities who themselves are responsible for their collection.
- The secondary data, on the other hand, are those which have already been collected by some other agency and which have already been processed.

Methods of collecting - Primary Data

1) Observation Method

- The investigator collects the requisite information personally through observation.
- The investigator meets the students in their hostels and collects necessary data after a personal study.

2) Questionnaire Method

- The investigator collects the requisite information by personally interviewing individuals, recording their answers in a structured questionnaire.

3) Mailed Questionnaire Method

- A set of questions relevant to subject of enquiry are mailed to a selected list of persons with a request to return them duly filled in.
- This method can only be used when the respondents are literate and can answer the questions in writing. The questions should be very clear without any ambiguity.

4) Telephone Interview

- This method of collecting information consists in contacting respondents on telephone itself.

Advantages of Primary Data collection methods

	Personal Interview with Questionnaire		Mailed Questionnaire Advantages		Telephone Interview
i)	Most flexible method of obtaining data	i)	Wider and more representative coverage possible at the same cost	i)	Wider and more representative coverage possible at the same cost
ii)	Identity of respondent known	ii)	No field Staff is required	ii)	No field Staff is required
iii)	Non response very low	iii)	Cost per questionnaire relatively low	iii)	Cost of response low
iv)	Supervision and control possible	iv)	No bias of interviewers	iv)	Quick way of obtaining data
		v)	Respondents can answer question directly		

Preferred Primary Data collection methods

Funds	Time	Type of Data	Personal Interview	Mail	Telephone
Restricted	Restricted	Few Items			X
Restricted	Restricted	Much Information	X		
Restricted	Ample	Few Items		X	X
Restricted	Ample	Much Information	X	X	
Ample	Restricted	Few Items	X		X
Ample	Restricted	Much Information	X		
Ample	Ample	Few Items	X		X
Ample	Ample	Much Information	X		

Methods of collecting - Secondary Data

1) Various publications of Central, State and local governments

- Statistical Abstract, India-Annual; Monthly Abstract of Statistics; Indian Agricultural Statistics, Ministry of Food and Agriculture; Index Number of Wholesale Prices in India Published by Ministry of Commerce and Industry; Reserve Bank of India Bulletin.

2) Various publications of foreign governments or of international bodies

- The international bodies like UNO, FAO, WHO, UNESCO, ILO, Statistical Year Book (Published by the Statistical Office of the United Nations), Yearbook of Labour Statistics (Published by ILO, Geneva).

3) journals of trade, commerce, economics, engineering etc

- Trade associations, Chambers of Commerce, Annual Report of the Chief Inspector of Mines in India, Indian Textile Bulletin issued by the Textile Commissioner, Bombay.

4) The other sources of secondary data

- books, magazines and newspapers, reports prepared by various universities, historical documents, diaries, letters, unpublished biographies and autobiographies.

Scrutiny of Secondary Data

The scrutiny should be made to assess the suitability, reliability, adequacy and accuracy of the data to be compiled and to be used for the proposed study.

- **Suitability :** The compiler should satisfy himself that the data contained in the publication will be suitable for his study. In particular, the conformity of the definitions, units of measurement and time frame should be checked. For example, one US gallon is different from one British gallon.
- **Reliability :** The reliability of the secondary data can be ascertained from the collecting agency, mode of collection and the time period of collection. For instance, secondary data collected by a voluntary agency with unskilled investigators :;re unlikely to be reliable.
- **Adequacy :** The source of data may be suitable and reliable but the data may not be, adequate for the proposed enquiry. The original data may cover a bigger or narrower geographical region or the data may not cover suitable periods. For instance, per capita income of Pakistan prior to 1971 is inadequate for reference during the subsequent periods as it became separated into two different countries with-considerable variation in standard of living.
- **Accuracy :** The user must be satisfied about the accuracy of the secondary data. The process of collecting raw data, the reproduction of processed data in the publication, the degree of accuracy desired and achieved should also be satisfactory and acceptable to the researcher.

Design of Questionnaires

'Questionnaire' is a commonly used and frequently abused tool for gathering a variety of data answered by an individual or group, with the aim of obtaining relevant data on the topic of research.

Types of questionnaires

(A) Structured questionnaires: They are prepared well in advance and not on the spot. The. form of questions may require responses which are either closed or open.

Close-ended questionnaires: Close-ended questionnaires are used when categorized data are required. They include a respondent can reply in a limited number of ways - 'yes', 'no', 'no-opinion', or an answer from a short list of possible responses. He/ she is asked to put a tick mark in a space provided on the answer sheet or is requested to underline a response.

Open-ended questionnaires: the respondent is asked to write a descriptive essay and express his/ her viewpoints or report on details and events, without restrictions

(B) Unstructured questionnaires: Flexibility is the chief advantage of the unstructured questionnaire. It is designed to obtain viewpoints, opinions, attitudes and to show relationships between various types of information which might escape notice under more mechanical types of interrogation. No predetermined responses are provided; instead, free responses are solicited.

Characteristics of a good questionnaire

1. **Purpose:** First, it must translate the objectives of an investigation into specific questions, the answers to which will provide the data necessary to test the hypotheses and explore the area defined by the objectives. Secondly, the questionnaire must motivate the respondents to communicate the required information.
2. **Language:** The language of a good questionnaire should be concise and directed towards producing uniformity of understanding among the respondents. The vocabulary should be simple and within the easy grasp of the least intelligent of the group under study. The syntax should be clear and straightforward. Vague phrases and expressions should be avoided. Technical expressions should be used only if the inquiry is directed to a select group which is well-versed in the technical language used. Proverbs and quotations should be avoided. Subjective words, such as 'bad', 'good', 'fair' and the like do not lend

Characteristics of a good questionnaire

3. **Frame of Reference:** Complex questions that require the respondent to go through several steps of reasoning before answering are undesirable and have often resulted in misleading information. A series of specific questions is needed so as to uncover degrees of intensity of feeling or conviction.

1. **Arrangement of Questions:** The arrangement or ordering of questions should receive special attention. It should appear logical to the respondents. The questions placed first in the questionnaire should be the easiest to answer. 'Interest-generating' questions should be asked at the beginning.

5. **Length of the Questionnaire:** A questionnaire should not be longer than necessary. The total number of questions must not be too large to tire or bore the respondents. If too many questions are asked and the respondent becomes tired, the questions at the end of the series may not be well answered.

6. **Form of Response:** The form in which the responses are recorded must be integrated with the form of the questions. There should be no hesitation in asking for responses in different forms in the same questionnaire.

Sampling and Sampling Designs

- The terminology "sampling" indicates the selection of a part of a group or an aggregate with a view to obtaining information about the whole.
- This aggregate or the totality of all members is known as **Population** although they need not be human beings.
- The selected part, which is used to ascertain the characteristics of the population is called **Sample.**
- The total number of members of the population and the number included in the sample are called **Population Size** and **Sample Size** respectively.
- A Researcher adopts a Technique to select the items of the Sample from the Population and that is called as Sampling Design.
- Sample design must be done before Data Collection.

Steps in Sample Design

Box 1 Sample Design Process

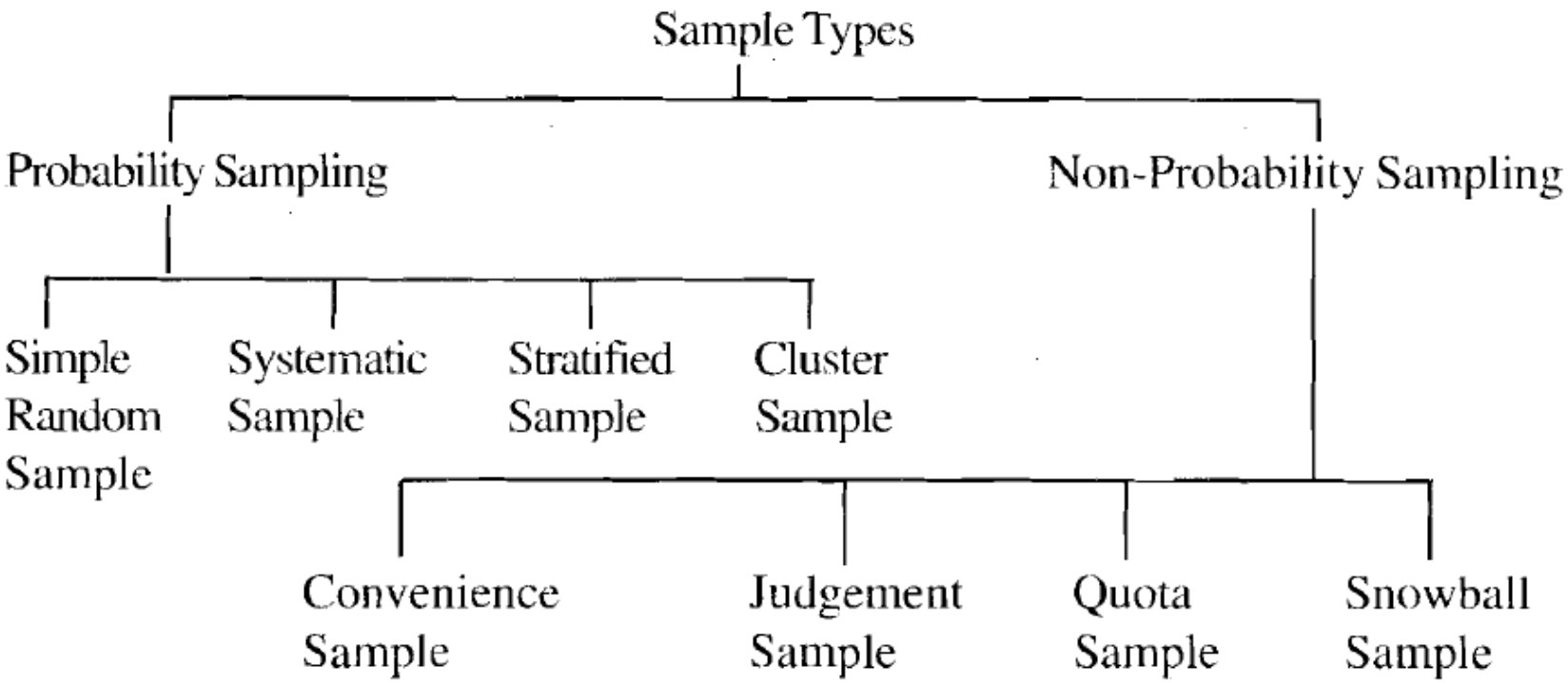
Step-1	Define the Population
Step- 2	Decide if it would be better to take a sample or a census
Step-3	If possible, construct or obtain an appropriate sampling frame
Step-4	Decide whether to use a probability sample or a non-probability sample
Step- 5	Select the sampling method

Source: John Boyce (www.mhhe.com/av/boyceze)

Factors Determining Sample Size

- The size of the universe
- The resource available
- The degree of accuracy or precision desired
- Homogeneity or heterogeneity of the universe
- Nature of study
- Methods of sampling adopted
- Nature of respondents

Sampling Methods:



Sampling Methods:

(1) Probability/ Random Sampling:

- Every member of the population has a equal chance of being selected. The members are free of bias. Error is more, less time and cost and less reliable information.

Types of Random Sampling:

(a) Simple Random Sampling: (Lottery Sampling)

The random sample entails that each and every individual in a population has an equal chance of being included in the sample and that the selection of one individual is in no way dependent upon the selection of another person.

Exp: In lottery draw, for example. if we have to select a sample of 25 students from a total of 600 students in a college, then we make separate slips of paper for 600 students and put them in a box and thoroughly mix them. After that, a person is asked to pick up one slip. Here. the probability of each of the student being selected in the sample is 11600. This procedure is continued till the sample size is acquired.

Sampling Methods:

- **(b) Systematic Sampling:** Designing a Systematic Random Sample is sometimes quite difficult and time consuming and therefore, Systematic Random Sample, like Simple Random Sample, also uses a list of all members of the population in its sampling frame. However, instead of using random numbers to select the sample elements. The researcher applies a skip interval to the list to produce a sample of the required size.

$$\text{Skip interval} = \frac{\text{number of elements in the population}}{\text{the required sample size}}$$

K = skip interval

N = Universe size

n = Sample size

$$K = \frac{N}{n}$$

- For example if we have to select a sample of 100 persons from a universe of 1000 population. then the skip is 10. In this case one number between 1 and 10 has to be selected. Suppose *5 is selected, then the first sample would be* 5^{th} *and* the next one 15^{th} , 25^{th} , 35^{th} , 45^{th} and so on. One of the advantages of this method is that it is more convenient than other methods and simple to design. Again, it is used with very large populations.

Sampling Methods:

- **Stratified Sampling:** In Stratified Random Sampling, the target population of N units is first divided into k subpopulations of N1, N2, N3 ....., Nk units. These populations are non-overlapping and together they comprise the whole population. So that N1 + N2 + ... + Nk = N respectively If the total sample size 'n' is to be drawn from the target population than n1 + n2 + ... + nk = n
- For example, A selected village]nay have households of SC(10%). ST *(5%), OBCs (45%).* Others (30%). Minority (10%). A village sample of 100 may constitute the households of various casts in the above proportional percentage so that the sample may contain all important characteristics of village population.

Sampling Methods:

- **(d) Cluster sampling:** Cluster sampling is a sampling technique used when natural groupings are evident in a statistical population. It is often used in marketing research. In this technique, the total population is divided into these known groups (or clusters) and a sample of the groups is selected. Briefly, the procedure for selecting a cluster sample is given below.
  - The population is divided into N groups, called clusters.
  - The researcher randomly selects n clusters to include in the sample.
  - The number of observations within each cluster is known:
  - M=M1 + M 2 + M 3 + ... +MN
  - Each element of the population can be assigned to one, and only one, cluster.

  As discussed above, in the cluster sampling method, the primary selecting unit is not a household, rather a natural cluster of households, viz., hamlets in villages, or, created clusters, viz., schools. malls, etc., may be decided. The first list of clusters may be selected using the SRS or the PPS sampling techniques. Then, from each selected cluster,. all units, or, some of the units, may be selected as per the required sample size using Stratified Random Sampling or the Systematic Random Sampling techniques.

Sampling Methods:

(2) Non Probability Sampling:

- Selection of sample in a non random order. It is opportunistic or purposive. It is based on convenience or the judgment of researcher.

Types of Non Probability Sampling:

(a)Convenience Sampling: The convenience sample is so called because it is relatively easy to obtain and contact. In this method the investigators are usually asked to select the people for the interview in accordance to the instructions from the researcher. The benefit of a convenience sample is that the interviewer **can usually get interviews done quickly** and cheaply. Convenience sampling is appropriate for exploratory research.

Sampling Methods:

(b) Quota Sampling: the population is first segmented into mutually exclusive sub-groups. Then judgment is used to select the subjects or units from each segment based on a specified proportion. For example, an interviewer may be told to sample 200 females and 300 males between the age of 45 and 60.

(c) Judgment sampling: A judgment sample is similar to that of convenience sample. In a judgment sample, the researcher selects samples that are believed to represent the population. The selection of samples is based on the knowledge of the population and the characteristics which the sample is to represent. It is less costly and very useful for forecasting.

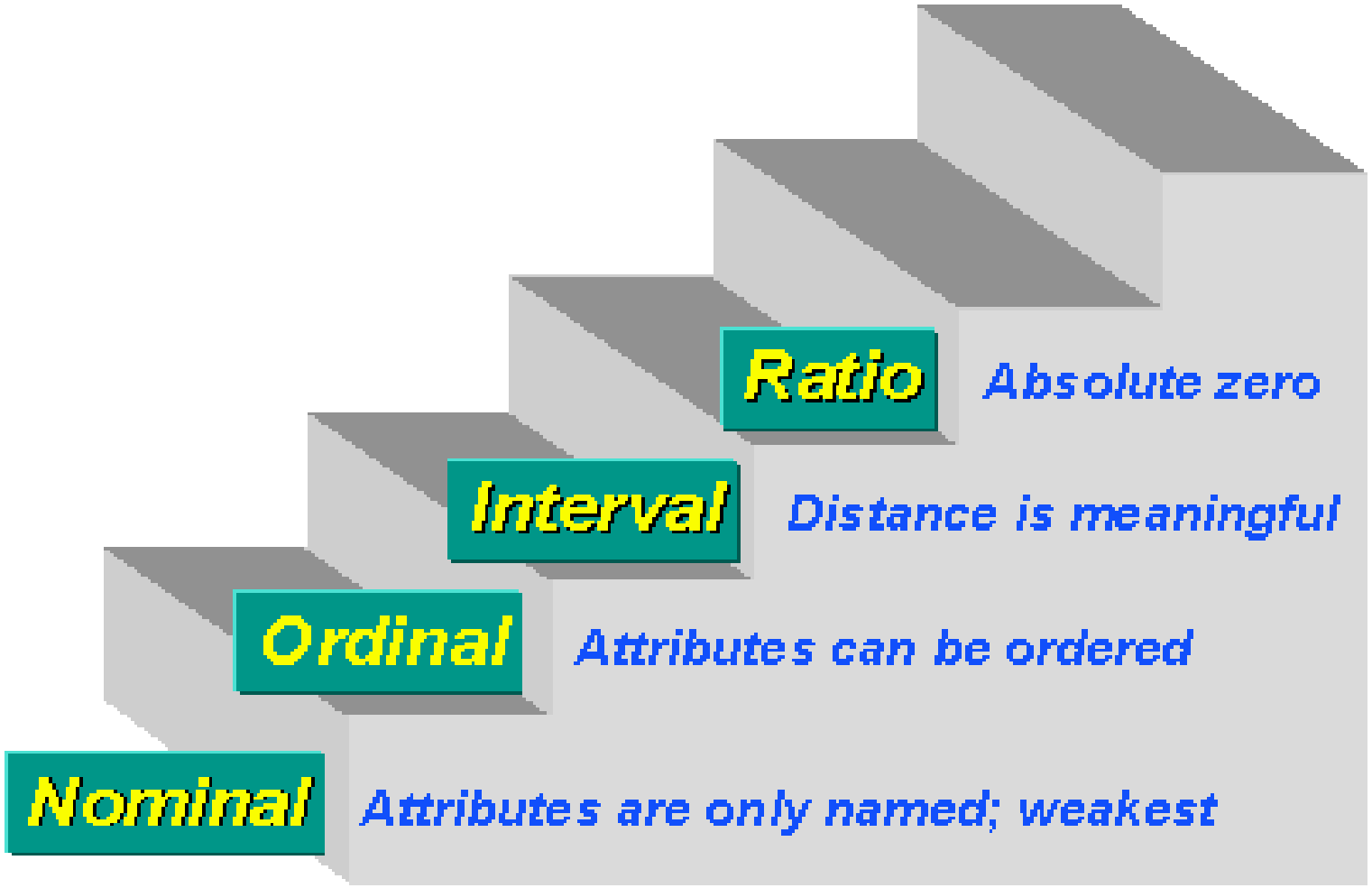
(d) Snowball Sample: This is one of the important types of non-probability sampling. In snowball sampling, the investigator encourages the respondents to give the names of other acquaintances and it continues growing in size and chains until the research purpose is achieved.

Measurement and Scaling Techniques

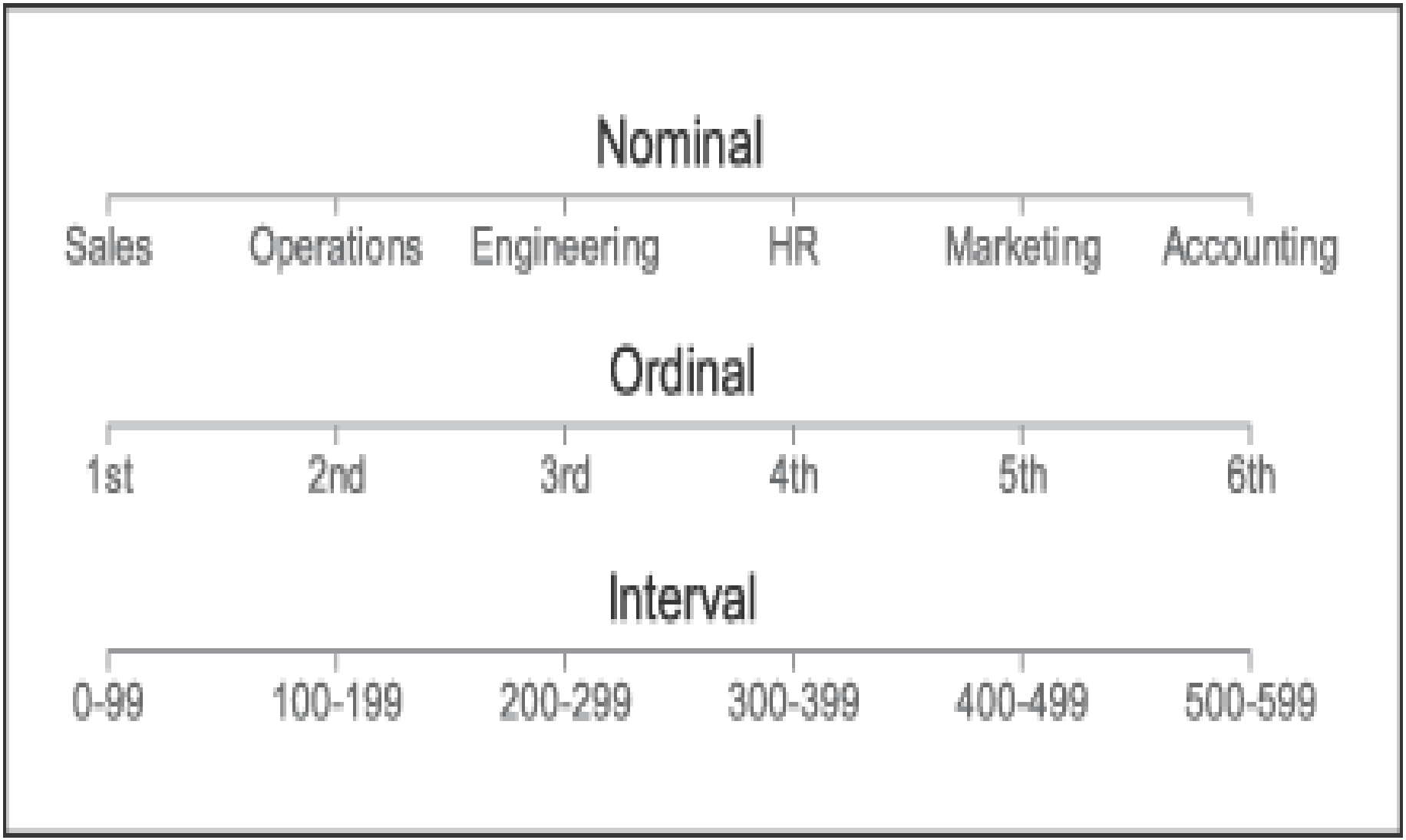
A ***scale*** provides the **unit of measurement** (such as categories, inches, seconds, etc.) and the **range** of possible values.

- The unit can be either **continuous** (i.e., allows partial units, or decimal, values) or **discrete** (only whole units), but the value of any particular unit has the *same meaning for all objects* that are assigned that value on the scale.
- A scale can either have an **infinite** range of possible values, meaning that a measure can take on any real-number value, or may be limited to **as few as two** possible values (if there is only one possible value, you have a **constant**, and you don't need a scale!).

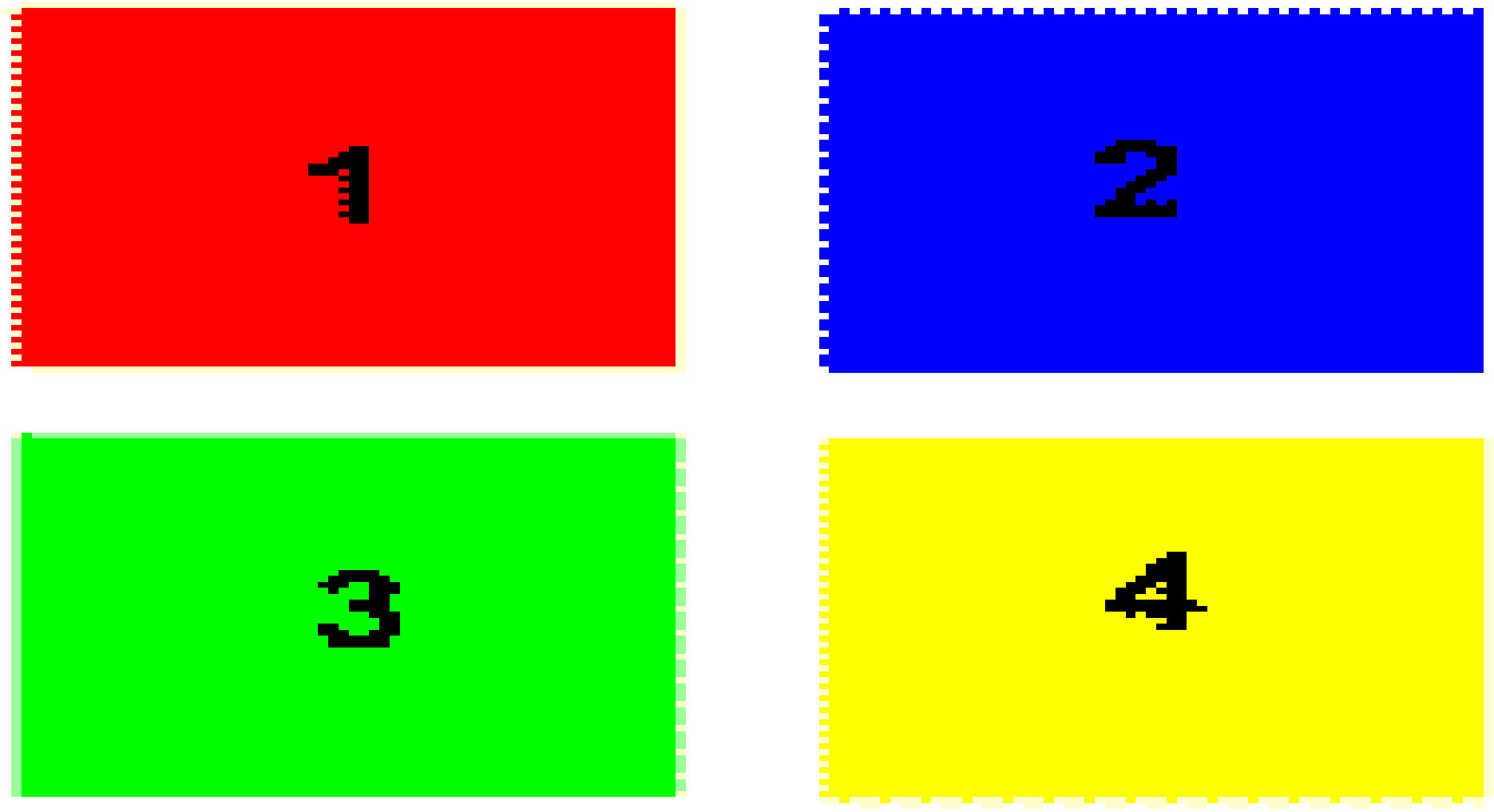
Different types of Scales



examples


Nominal
Sales
Operations
Engineering
HR
Marketing
Accounting
Ordinal
1st
2nd
3rd
4th
5th
6th
Interval
0-99
100-199
200-299
300-399
400-499
500-599

Nominal scale

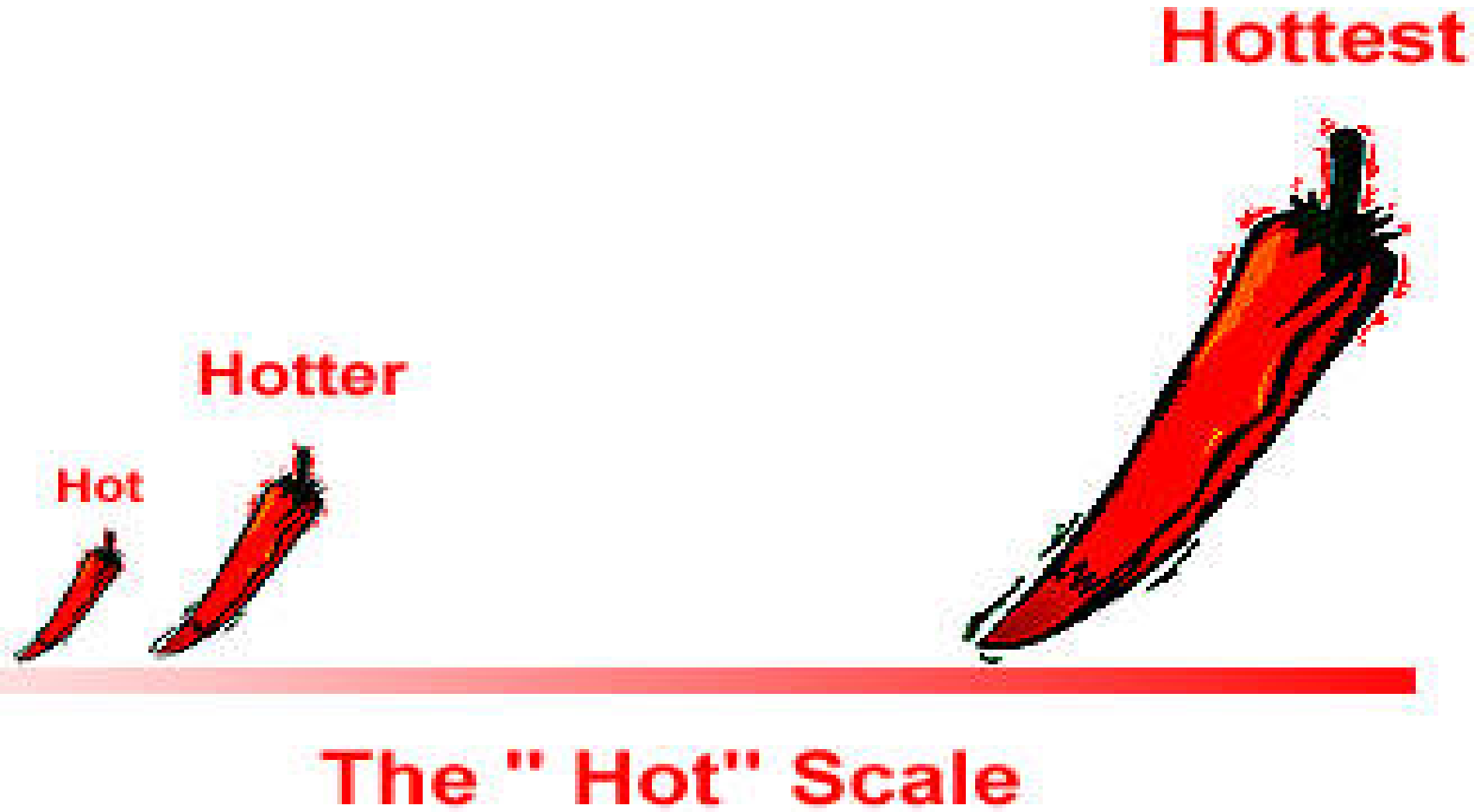


Nominal scale

- With nominal scales, information is collected on a variable that naturally or by design can be grouped into two or more categories that are mutually exclusive and collectively exhaustive.
- If we divide students on basis of their mother tongue by assigning number like:
  - **English-1**
  - **Hindi-2**
  - **Other-3**

  Then every student is expected to write their language by selecting any one of the above digit.
- The counting of members/elements in each group is the only possible arithmetic operation when a nominal scale is employed.

Ordinal scale



Ordinal scale

Ordinal Data

Point	**Airports** international national regional	**Oil well production** high medium low	**Populated places** large medium small
Line	**Roads** expressway major local	**Drainage** river stream creek	**Boundaries** international provincial county
Area	**Soil quality** good fair poor	**Cost of living** high medium low	**Industrial regions** major minor

Ordinal scale

- Ordinal scales include the characteristics of the nominal scale plus an indication of order.
- Ordinal data require conformity to a logical postulate which states: if ‘a’ is greater than ‘b’ and ‘b’ is greater than ‘c’ then ‘a’ is greater than ‘c’.
- The use of an ordinal scale implies a statement of greater than or less than without stating how much greater or less.
- Also some other descriptors like highly agree, agree, somehow agree or most successful, successful etc can be used.

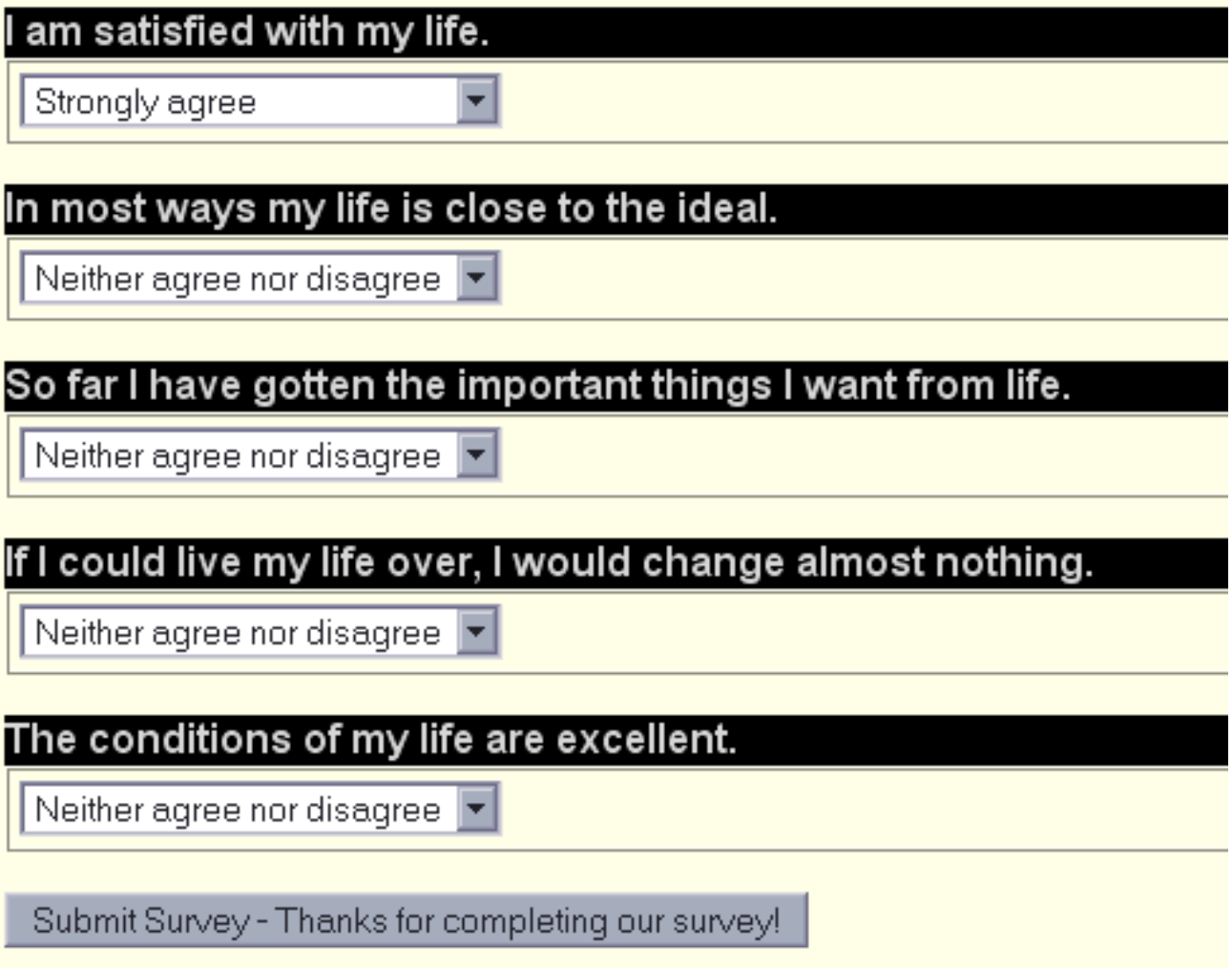
Interval scale

- The interval scale have the power of nominal and ordinal data plus one additional strength.
- They incorporate the concept of equality of interval .
- It is a scale that not only arranges objects or alternatives according to their magnitudes but also distinguishes this ordered arrangement in units of equal intervals

Interval scale

- In **interval** measurement the distance between attributes *does* have meaning.
- For example, when we measure temperature (in Fahrenheit), the distance from 30-40 is same as distance from 70-80.
- The interval between values is interpretable. Because of this, it makes sense to compute an average of an interval variable, where it doesn't make sense to do so for ordinal scales.
- But note that in interval measurement ratios don't make any sense - 80 degrees is not twice as hot as 40 degrees

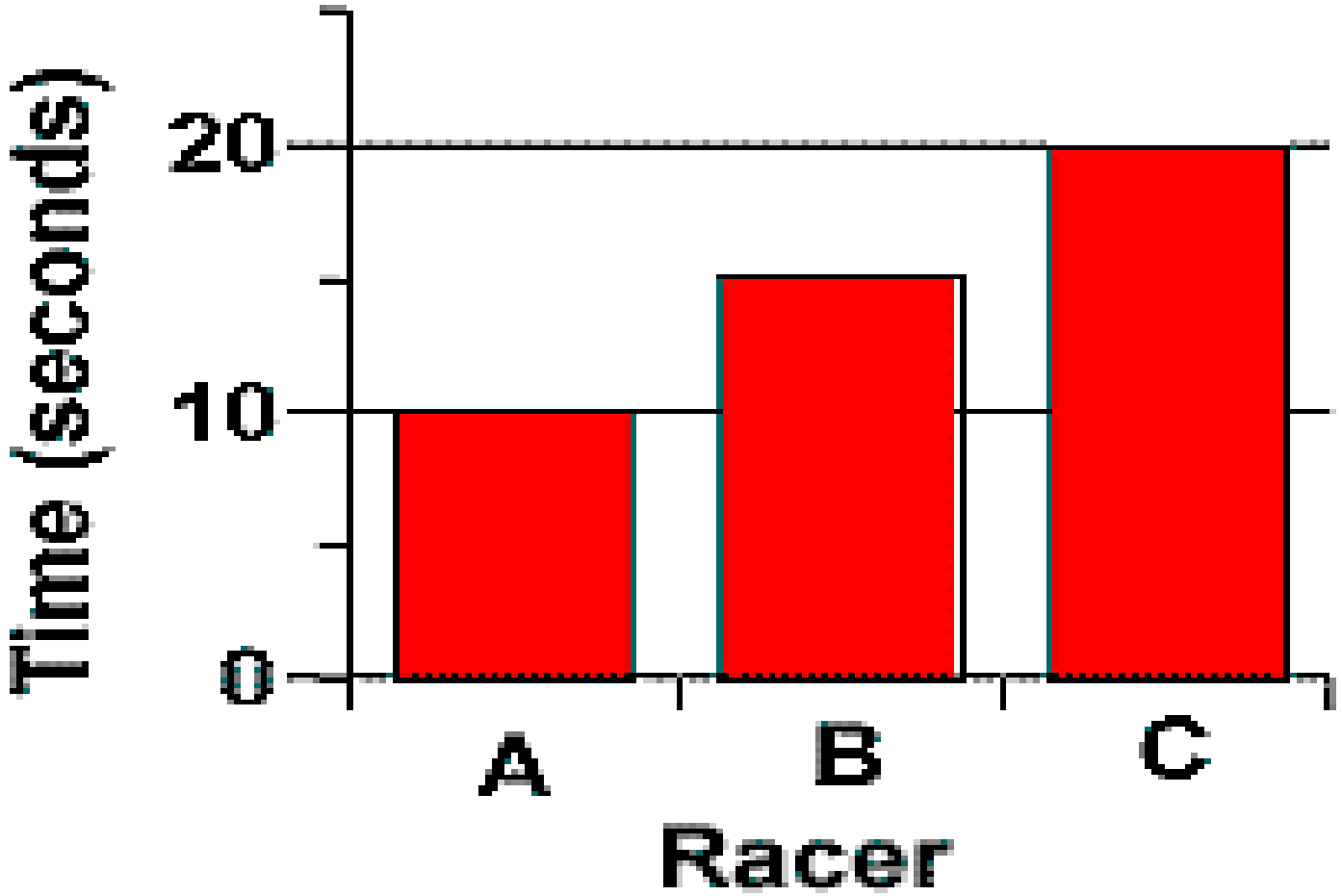
Interval scale



Ratio scale

- Ratio scales are similar to interval scales.
- A ratio scale allows you to compare differences between numbers.
- For example, if you measured the time it takes 3 people to run a race, their times may be 10 seconds (Racer A), 15 seconds (Racer B) and 20 seconds (Racer C). You can say with accuracy, that it took Racer C twice as long as Racer A. Unlike the interval scale, the ratio scale has a true zero value.

Ratio scale



Ratio scale

- Finally, in **ratio** measurement there is always an absolute zero that is meaningful.
- This means that you can construct a meaningful fraction (or ratio) with a ratio variable. Weight is a ratio variable.
- In applied social research most "count" variables are ratio, for example, the number of clients in past six months. Why?
- Because you can have zero clients and because it is meaningful to say that "...we had twice as many clients in the past six months as we did in the previous six months."

Some special types of Rating Scales

1. Simple attitude scales
2. Semantic differential scales
3. Likert scale
4. Graphic rating scales
5. Sorting
6. Staple scale
7. Numerical/multiple rating list scales
8. Constant sum scales
9. Ranking scales
10. Cumulative scales

Measurement error

- Measurement error or measurement differences occur due to non attainment of complete control over the variables under study.
- There are four measure sources of error:

  - **The respondent**
  - **Situational factor**
  - **The measurer**
  - **The data collection instrument**

Data Processing

- VALIDITY
  - **Content**
  - **Criterion related**
  - **construct**
- TEST OF RELIABILITY
  - **Stability**
  - **Equivalence**
  - **Internal consistency**
- TEST OF PRACTICALITY
  - **Economy**
  - **Convenience**
  - **interpretability**

Test of VALIDITY

- The ability of a scale or measuring instrument to measure what it is intended to measure is tested in test of validity.
- Not everything that can be counted counts, and not everything that counts can be counted.

By Albert Einstein

Test of Reliability

- A measure is reliable to the degree that it supplies consistent results. Reliability is a necessary contributor to validity but is not a sufficient condition for validity.
- The test of reliability implies the degree to which measures are free from error and therefore yield consistent results.
- When the outcome of the measuring process is reproducible, the measuring instrument is reliable.

Equivalence

- In the equivalence form method two alternative instruments are designed to be as equivalent as possible. Each of the two measurement scales is administered to the same group of subject. If there is high correlation between the tow forms , the researcher concludes that the scale is reliable

Stability

- A measure is said to possess stability if you can secure consistent result with repeated measurements of the same person with the same instrument.
- An observation procedure is stable if it gives the same reading on a particular person when repeated one or more times

Internal consistency

- A third approach to reliability uses only one administration of an instrument or test to assess the internal consistency or homogeneity among the items.
- The split half technique can be used when the measuring tool has many similar questions or statements to which the participant can respond.
- The instrument is administered and the results are separated by item into even and odd numbers or into randomly selected halves
- When the two halves are correlated, if the results of the correlation are high, the instrument is said to have high reliability among the items.

Test of Practicality

- Practicality has been defined as economy, convenience and interpretability.
- The scientific requirements of a project call for the measurement process to be reliable and valid, while the operational requirements call for it to be practical.

Economy

- Some trade off usually occurs between the ideal research project and the budget.
- Data are not free, and instrument length is one area where economic pressures dominate.
- More items give more reliability, but in the interest of limiting the interview or observation time, we hold down the number of measurement questions.
- The choice of data collection method is also often dictated by economic factors.

Convenience

- A measuring device passes the convenience test if it is easy to administer
- A questionnaire or a measurement scale with a set of detailed but clear instructions, with examples, is easier to complete correctly than one that lacks these features.

Interpretability

- This aspects of practicality is relevant when persons other than the test designers must interpret the results.
- It is usually, but not exclusively an issue with standardized test.
- In such cases the designer of the data collection instrument provides several key pieces of information to make interpretation possible.

Data analysis

- Averages
- Variances
- Proportions
- Std. Deviation
- Hypothesis testing
- Regression
- Correlation
- Factor analysis
- Cluster analysis
- Multivariate analysis
- Discriminant analysis